



Blockchain-Based Traceability Method - A Review

David Wong You King^{1,2}, Muhammad Arif Riza¹, Liew Kok Leong^{1,2(✉)},
Ummul Hanan Mohamad¹, Rabiah Abdul Kadir¹, Mohammad Fairus Zulkifli¹,
and Mohammad Nazir Ahmad¹

¹ Institute of Visual Informatics, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor,
Malaysia

p119039@siswa.ukm.edu.my

² 22-08, Level 22, Lingkaran TRX, Tun Razak Exchange, Menara Exchange106,
55188 Kuala Lumpur, Malaysia

Abstract. Effective output optimization in various industries requires precise monitoring of all processes and data involved. To facilitate further optimization, information must be easily accessible and traceable. While data traceability can be achieved using various technologies, some methods may have limitations, such as a lack of real-time monitoring or difficulty integrating with existing systems. In recent years, the emergence of blockchain technology has enabled new possibilities for data traceability. Blockchains are composed of blocks of hashed data that are linked to earlier blocks, providing a reliable means of tracing data. This paper explores the fundamental concepts of blockchain technology and its features, highlighting the limitations of traditional traceability systems and the advantages offered by blockchain. Then, this paper reviews existing literatures to provide a comprehensive overview of blockchain-based traceability methods, offering insights into their applications, benefits, challenges, and future prospects.

Keywords: Consortium · Decentralized · Ethereum · Industrial Process · Smart Contracts

1 Introduction

Data traceability is crucial for optimizing the quality of a product or a service. For example, in food industries, product traceability is important to help identify the processes that each product undergo. Meanwhile, agricultural industries require crop traceability to ensure optimized quality in large batches of crops. To add, service or tourism industries need data traceability to ensure proper transactions and the security of customers' data (Nyalety, et al. 2019). However, with the existence of many layers of the whole process, traceability of data can be very difficult to manage. Due to this, many large industries with large production capabilities experience difficulties in tracing each of their products for quality checks.

Traditional traceability methods, such as barcodes and quick response codes (QR), are cost-effective and easy to implement (Qian et al. 2012; Tarjan, et al. 2014). However,

they lacked real-time monitoring capabilities and do not provide enough information to conduct detailed analysis and optimize the traceability process.

Artificial intelligence (AI) technology has emerged as a potential solution to enhance data traceability. AI can provide early warnings of potential errors in the traceability process, allowing organizations to proactively address any issues beforehand. For example, AI algorithms can analyze data from multiple sources, including sensors and tracking systems, to identify anomalies and deviations from expected patterns. However, AI technology is complex, therefore it requires highly skilled workers to operate and maintain it (Zhang and Lu 2021).

The Internet of Things (IoT) is another technology that has the potential to enhance data traceability. IoT involves connecting devices and sensors to the internet, allowing for remote monitoring and real-time data collection. When integrated with traceability systems, IoT can provide valuable insights into the journey of a product or service, helping organizations to optimize their processes and improve the quality of their offerings. However, implementing IoT can be costly and the devices can be vulnerable to hacking, which could compromise the traceability of the product (Cui, Chen, et al. 2019).

Differing from IoT, blockchain (BC) technology is based on its implementation. It is a method that is less susceptible to external attacks due to its decentralized nature. In general, blockchains are a chain of information that are stored in the form of hashes within a block. Information blocks will be created, and a newer block will be generated following the previous block for new information. This creates a link of information blocks that can easily be accessed and traced. As more blocks are created in the chain, the system is considered more stable and less susceptible to data attacks (Jiang et al. 2022).

Since blockchain technology has emerged as a promising solution to address the data traceability challenges, this review paper analyzes existing methods and frameworks for establishing traceability, covering industries such as food, pharmaceuticals, luxury goods, and electronics. It evaluates the benefits and challenges associated with implementing blockchain-based traceability systems, addressing scalability, interoperability, and regulatory considerations. Additionally, the paper identifies emerging trends and future research directions in the field, emphasizing the need for interdisciplinary collaboration, standardization efforts, and advanced analytics. Overall, this review serves as a valuable resource for researchers, practitioners, and policymakers interested in harnessing blockchain technology for transparent and accountable traceability.

2 Blockchain Traceability Methods

Applications of BC in traceability are still at infancy stage and under continuous improvement. There are several methods discussed from other studies as shown in Table 1. Each of the methods was used in conjunction with data traceability for products and industries.

Data traceability in IoT involves the tracing of data from various devices that are connected to the internet. All kinds of data are continuously generated from different fields and industries. A large amount of data will require sufficient security measures and handling. However, due to the enormous amount of data that stems from various devices, data traceability can be a challenge if conventional methods are used. Hence, Qiao

Table 1. Various methods for blockchain traceability methods.

Blockchain Traceability Technique	Application Domain	Mechanism	Advantages	Disadvantages	Design/run time	Data standardization	References
Consortium blockchain	Dynamic data tracing in IoT	<ul style="list-style-type: none"> - Verification nodes list - Authorizing some of the trusted nodes based on the boundary conditions 	<ul style="list-style-type: none"> - Consortium blockchain can set the openness to the public according to the application scenario - Auditing of dynamic data by multiple agencies in a specific background 	<ul style="list-style-type: none"> - Only approved nodes to join in is more suitable for the management of dynamic data 	High run time	Possible but requires extensive mechanism modification	Qiao et al. 2018
Custom mathematical model based on Food Quality Index (FQI) algorithm	Food quality traceability	<ul style="list-style-type: none"> - Mathematical modeling is devised for the calculation of the food quality index considering the storage time as the selected quality parameter 	<ul style="list-style-type: none"> - Suitable for traceability in the food industry due to its nature of implementation 	<ul style="list-style-type: none"> - Model is tailored to food quality traceability only. Substantial mathematical model modification is necessary for other applications 	High design time for building the modeling	Not possible	George et al. 2019

(continued)

Table 1. (continued)

Blockchain Traceability Technique	Application Domain	Mechanism	Advantages	Disadvantages	Design/run time	Data standardization	References
Hyperledger Fabric (consortium blockchain)	Food safety risk traceability	Multiple components at each part of the system: Orderer, Client, Endorser, Committer	<ul style="list-style-type: none"> - Visualization methods to intuitively show risks and help the traceability analysis of food - Good tamper security 	<ul style="list-style-type: none"> - Slow processing speed and scalability of generating blocks - The visual analysis data quality may be affected by sampling errors 	High run time due to slow processing speeds	Possible but difficult to achieve	Hao et al., 2020
Control policy model based on smart contract	Medication Traceability	<ul style="list-style-type: none"> - Byzantine fault tolerance (PBFT) modeling for score-keeping mechanism, allow the consensus algorithm to optimize upon facing dishonest nodes 	<ul style="list-style-type: none"> - Eliminates the need for centralized institutions and third-party organizations - Provides a full record of the medication circulation process - High level of security and privacy protection 	<ul style="list-style-type: none"> - The study was only reported on a small scale on a single PC rather than clustering deployment - All information is stored on the blockchain as proposed leading to poor memory utilization 	Low run time	Not possible	Zhu et al., 2020

(continued)

Table 1. (continued)

Blockchain Traceability Technique	Application Domain	Mechanism	Advantages	Disadvantages	Design/run time	Data standardization	References
Ethereum smart contracts	COVID-19 tracking from external sources	-The system connects decentralized applications (DApps), dashboards, smart contracts, oracles, and web feed sources within the same decentralized Ethereum network	<ul style="list-style-type: none"> - Receives data from various web feed sources - On one of the well-known blockchains – Ethereum 	<ul style="list-style-type: none"> - Uncertainty of implementations and operations costs due to the fluctuation of Ethereum gas fees for transactions - The blockchain network traffic becomes bulky as the number of transactions increases every day - Restriction on the block size and time interval used to create a new block 	<ul style="list-style-type: none"> -Low design time -Low run time 	Not possible	Marbough et al. 2020
Raft consensus algorithm	Personal Information Registration	-Each same nodes have multiple states. Can switch states/roles	The entities involved in the transaction can obtain sufficient transaction-related evidence to resolve possible future disputes after the execution of the transaction	- Susceptible to info leaks	High design time	Possible but difficult to achieve	Jiang et al. 2020

(continued)

Table 1. (continued)

Blockchain Traceability Technique	Application Domain	Mechanism	Advantages	Disadvantages	Design/run time	Data standardization	References
Stochastic batch dispersion model and Negative binomial distribution	Food supply chain traceability	<ul style="list-style-type: none"> - Reduction of mixing of data in batches - After attaining traceability, 'batch dispersion problem' data is integrated with Blockchain to achieve transparency. The second part of the problem is to ensure the safety of this data inside the 'distributed ledger' 	<ul style="list-style-type: none"> - The stochastic model and the optimum batch size give an insight into the composition of each product which proves more the amount of data is fed into the Blockchain, the safer it becomes 	<ul style="list-style-type: none"> - Susceptible to attack at early implementation - 5-level model dispersion was created specifically for food production. Extensive modification of the model may be required and tested for other use 	High design time	Not possible due to modeling is tailored only specifically to the use case	Maity et al., 2021
Traceability system based on Ethereum SHA-3 algorithm	Smart farming	<ul style="list-style-type: none"> -Entry information, Inquiry, and verification with the algorithms 	<ul style="list-style-type: none"> -Distributed ledger using blockchain technology -Complete Decentralization 	<ul style="list-style-type: none"> The system is implemented together with QR code technology. Thus share the disadvantages of QR code technology for traceability -Data standardization can be a challenge 	Low design time	Possible but difficult to achieve	Liao et al. 2019

et al. (2018) proposed the traceability of data using consortium blockchain. Consortium blockchains are accessible to the appointed members of an organization. These members can set the read-write access permission and assign additional rules to the blockchain. In the study, the blockchain mechanism follows a verification nodes list which is formulated based on a set of boundary conditions. Each node possesses its own credibility based on its behavior. The credibility of the nodes was maintained based on how well they served other nodes in data verification. A node's credibility will be negatively impacted if it is inactive or denied service for a period of time. The node with low credibility will be removed from the list specified by the governing agency.

The proposed method by Qiao et al. (2018) allowed the formation of a high-quality data chain in terms of a node of trustfulness and security. The formation of a data blockchain is achievable due to the boundary conditions applied to the consensus nodes. The boundary conditions are formulated to allow nodes with high credibility to the main data chain. Any foreign, dishonest, or attacking nodes will have lower credibility scores and will most unlikely to be accepted in the main chain. However, the boundary conditions for the consensus nodes do not eliminate the possibility of attacking nodes gaining acceptance into the main chain. Hence, the proposed data traceability method is still risky due to dishonest nodes having a chance to infiltrate the main chain. The node verification could lead to a high run time of the system and data standardization from mixed parameters can also be difficult.

In the food industry, data traceability with blockchain technology has been proposed (George, et al. 2019) with the aid of the Food Quality Index (FQI). Data traceability in the food industry has been implemented with different technologies throughout the years. The tracking of various information in food production such as product tracking, raw material information, and logistics allows continuous improvement of the overall operations. Barcodes and QR codes have been implemented to allow data traceability for products and materials. However, this technology can be difficult to implement, especially for large production-scale industries. It requires a carefully structured algorithm build that leads to high design time especially if data standardization is required. Therefore, with blockchain, it is possible to implement within an industry for data traceability due to its long-term effectiveness and scaling. When information is stored in a blockchain, its accessibility is more efficient compared to centralized storage.

The study proposed by Hao et al. (2020) was similar to Qiao et al. (2018) in terms of the boundary condition implemented. It was a model centered around parameters such as the shelf life, nutritional value, and storage time of the product. These data are collected via sensors and manually keyed in by farmers. The data is recorded and processed within the blockchain to determine its food safety. The determination of food safety is based on a food index constructed from prescribed standards made by the regulatory authority. A curve estimation based on shelf life versus weightage was made. Extremum points are derived based on the curve estimation and serve as the range of values determining whether the food is safe for consumption.

The proposed blockchain traceability method (George et al., 2019) based on the food index has potential use in the food and beverage (F&B) industries. The boundary conditions prepared based on the shelf life of food allow the construction of the food index. Since the food index can be different for each type of food, the method allows

for additional food indexes to be updated within the blockchain boundary conditions. However, this method can only be used for food products. Its workability is not explained for other kinds of production such as crops, electrical components, data services, etc. A different boundary condition with different algorithms will be required for other industries.

Blockchain technology is constantly being improved and combined with other technology for data traceability. A combination of blockchain technology and visualization technology has also been proposed. This blockchain traceability method was built around Hyperledger Fabric technology. This technology enables flexibility in terms of the number of parties involved in the development of a blockchain platform. The main component of Hyperledger Fabric is (Hao et al. 2020):

- a) Orderer – Receives all transactions throughout the main designated network before packing the data into blocks.
- b) Client – The user that is utilizing a software development kit (SDK) to access and request transactions from the Hyperledger network.
- c) Endorser – These nodes endorse clients that requested a transaction initiation. A client must obtain enough endorsements from the endorser node to fully initiate a transaction.
- d) Committer – Receives the packaged blocks from Orderers. These nodes verify the block's transaction validity before updating the ledger.

The proposed method in the mentioned study utilized several layers in its framework. A business layer serves as a contact point between a human user and a computer. The operation of uploading data and visualizing them through the display is the primary purpose of this layer. Users can write or update their smart contracts through the Application Programming Interface (API) that can be accessed through this layer. The communications layer of the framework houses the main protocols of the P2P network and network structures. Synchronization data between all nodes occur in the communications layer via the Gossip data communication protocol. The communications layer is a bridge between the business layer and the database layer. The stored data ledgers are kept within the database layer in the form of data blocks. The sophisticated framework with hyper ledger technology allows visualization mapping of data. The visualization is useful for mapping out data in terms of food quality to assist in the prediction of shelf life or identification of defects. However, the framework suffers from the speed of data processing, high program run time and block generation scaling. The added visualization of data suffers from extended delays as the number of data blocks increases.

The utilization of blockchain in the pharmaceutical and medical industry can be potentially beneficial. For medicinal supplies, it is required by the Global traceability standards that all items are assigned a unique identification number. The identification number aids in tracing the movement of pharmaceutical products throughout the supply chain. Zhu et al. (2020) proposed a traceability method with the aid of blockchain infrastructure. The pharmaceutical products will be assigned their ID number after manufacturing and this number will be recorded into the blockchain. Wholesalers and retailers serve as an intermediate node that can update the status of the medication on the blockchain (its location within the supply chain) via a peer-to-peer (P2P) network.

The system proposed by Zhu et al. (2020) consisted of a smart contract, a blockchain network, and a web client for the user. Smart contract coded using Python contains a mechanism of practical Byzantine Fault Tolerance (PBFT) for establishing a consensus among participating nodes. In the PBFT mechanism, the amount of network traffic can increase quickly due to nodes attempting to reach a consensus. This can lead to a slower time for information processing and transfer. This is remedied with the introduction of a score-keeping system. This system aids the nodes to reach consensus quickly before the network traffic is congested. Each node can receive a form of reward or punishment depending on the process of reaching a consensus. A dishonest node (carrying modified/tampered/forged information) from a set of nodes would cause the set to be deducted of score points. A set of nodes with low score points will have a lower probability to be selected for the consensus operation process. This ensures that the set of nodes with higher scores are nodes that can be trusted the most.

Network traffic will then be reduced since the low-scored nodes that can hinder the consensus process have less probability to participate. The system does show its potential as an organized BC-based data traceability system, however, the overall system is good for one form of data at a time. In the mentioned study, only the medication ID forms the hash before being participated throughout the system of nodes for validation. The scoring system would collapse if different forms or data parameters were to be validated. The system would suffer from a data standardization issue due to the nodes' inability to be given proper scoring or the potential in receiving penalties (Zhu, et al. 2020).

Besides the creation of an organization's own smart contracts and blockchain systems, it is also possible for data traceability to utilize public commercial blockchain networks for data traceability. It can be desirable to utilize an established and well-known blockchain network for data traceability for stability purposes. Marbough et al. (2020) proposed a data traceability method during the COVID outbreak in 2019. Resource deployments are a challenge, especially in the mass mobilization of vaccines and medical equipment. The proposed data traceability method involves the use of Ethereum blockchain network. The public blockchain network eliminates the need for internal, intermediary nodes for validation. Therefore, data traceability and validation can be achieved at less cost. In addition, the Ethereum blockchain also allows multiple smart contracts to be implemented for certain desired purposes.

There were three Ethereum smart contracts created by Marbough et al. (2020) for tracking of COVID-19-related data; a registration contract that handles the web information of participating stakeholders and web sources, a reputation contract that assigns scores based on the quality and trustfulness of web sources and an aggregator smart contract tasked with the retrieval of the latest information updates from nodes. The system allows for multiple information and parameters to be tracked within the blockchain network. Unlike privately developed blockchain networks, public network like Ethereum requires gas fees for each successful data transaction and smart contract code execution. More complex smart contracts with additional codes may lead to an increase in costs. Furthermore, with the price movement of Ethereum coins, its gas fee will be greatly impacted leading to difficulties in operational costing calculation.

For sensitive information traceability, a method was proposed by Jiang et al. (2020). The system was centered on a service provider, certification center and the user whose

information is to be transacted. For added security purposes, the system does not involve users to immediately upload their personal information to the certification center (can be government bodies, employers, etc.). The user is required to register and generate an identity in the system before the certification provides the new user with a unique identification number. The user then needs to encrypt his personal information with the provided identification number before sending the data to the service provider. The service provider then generates a random symmetric record key and encrypts the information package (that includes all the sensitive personal information).

The certification center (original information requestor) will decrypt the record key with the information package and sends the user the same record key for safekeeping. The certification center will check the record key and identification number. If both number matches and no tempering on all the backups of the package transactions, the certification center will release the data key to the service provider. This technique demonstrated a secure method for tracing sensitive data. However, it can be challenging in terms of scalability due to the numerous amounts of encryption involved. It also does not allow ease of data access for each node which can be difficult for organizations with multiple intermediary bodies serving as nodes in between.

A batch of information can be traced within the blockchain network with the proper mathematical modeling. Within the overall production process, various items and raw materials need to be tracked at once (within one transaction) and this can be challenging to achieve with minimal use of smart contracts and basic mathematical models of data traceability. One needs to tailor specific mathematical modeling depending on how many steps or phases are involved in the processing of several different raw materials into their respective end products. Most products nowadays average on undergoing 4–6 major processes before a raw material reaches its commercial form. A stochastic batch dispersion model was proposed by Maity et al. (2021) to achieve data traceability for food processing optimization. The batch processing of products makes it suitable for use with mathematical models that minimize the sum of links between the packaged product and its respective raw materials. This way, the inconsistency of data due to completed products from external sources can be minimized. The stochastic batch dispersion model considers 5 levels of process within the production supply chain. The individual levels represent the state of the product; therefore level 1 is the raw material, and level 5 is the finished product. There are two stages of the mathematical model with the first being the determination of fixed proportions of the finished raw materials of the product. The second stage of the model assist in which raw materials are required to be outsourced to fulfill any remaining customer demand. Secondary stage acts are the corrective measure for the mathematical model to ensure product traceability and the correct proportion of raw material used.

Blockchain systems can be integrated with sensors for quality inspection and data traceability, and this has been proposed (Liao and Xu 2019). The Ethereum blockchain is used as the underlying architecture for the system and is the data entry point. The information regarding the product is stored within the public blockchain and can be retrieved via a QR code. This system architecture shares a similarity with other studies (Marbouh et al. 2020) that involve separate layers of traceability system via the business layer, data layer, and user interfacing layer. Although the system can have data standardization due

to having the QR code as the form of information retrieval technique, the efficiency of the QR code falls behind newer tracking technologies such as VR/AR, IoT, Big Data, and AI.

3 Category of Blockchain Traceability

Based on the literature, there are at least three different categories of BC traceability methods. These methods vary greatly in terms of implementation, algorithm, and system structures.

3.1 Consortium Centric Blockchains

A consortium blockchain is one of the methods in BC-based traceability used. This technique has been utilized by Qiao et al. (2018). This technique involves several specialized nodes or trusted nodes for the verification of data chains. These networks of nodes are governed by personnel who is the member of an organization (alliance). The personnel can access the nodes through the agency gateway. The read and write access permissions can be granted based on the preferences of the alliance. The data in the blockchain can be made to be either private or public to all alliances. The system can be reliable and flexible due to the full control of alliance members and their transparency, but consortium-based blockchains can lead to slow processes due to the nodes waiting for the approval of other nodes. The distrustful nodes can still be chosen by other nodes to be accepted into the main chain. It is possible that consortium-centric blockchains on their own can suffer from reliability issues.

3.2 Smart Contracts

Blockchain networks such as Ethereum rely on proof of stake for node verification. Ethereum possesses smart contracts that can be utilized in data traceability. Smart contracts are software agents that are inserted in blockchain networks. Figure 1 is an example of a smart contract-based blockchain system operation. These agents enable the automatic verification of transactions with no interference from third-party entities. This increases the transaction approval efficiency and transaction processing time. In a work proposed by Marbough et al. (2020), an Ethereum smart contract was utilized for data traceability of COVID-19 cases. The study aims to solve the data traceability issues regarding medical supply chains and cases of COVID-19. These data can be overwhelming and lack security for a standard centralized data storage to handle. Therefore, the proposed work considered using three smart contracts within an Ethereum blockchain ecosystem to aid in data traceability. Smart contracts can be beneficial since it is open to flexibility for the user to add additional parameters to suit the industry traceability needs. It is easy to apply smart contracts to existing public domain networks such as Ethereum. However, the gas prices (fee) for transaction processing are inconsistent as the value of Ethereum fluctuates depending on the cryptocurrency market. Estimation of cost can be challenging due to the ups and downs of the gas fee pricing.

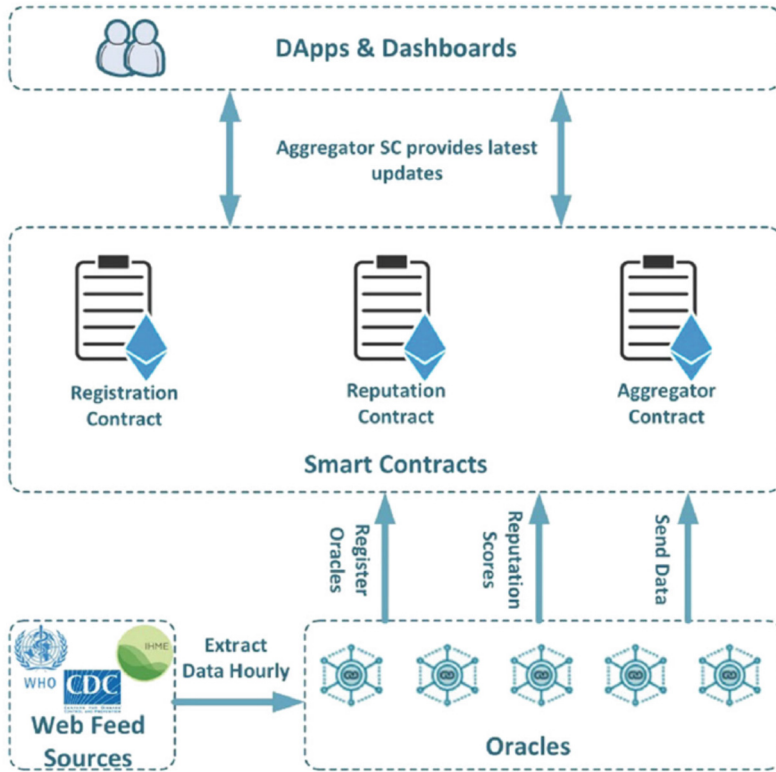


Fig. 1. A smart contract assisted blockchain data traceability process for COVID-19 tracking (Dounia, et al. 2020).

3.3 Custom Algorithms and Mathematical Modeling

Another method with data traceability with BC is the modification of custom algorithms. This requires detailed programming with a strong emphasis on the mathematical model built around the data traceability process and its elements. For instance, the Byzantine fault tolerance algorithm can enhance the blockchain process for nodes to achieve consensus (Zhu et al. 2020). This means that the algorithm can enhance other blockchain methods if well-coded. However, a strict and complex rule must be specified for each item or type of information. As more types of different information exist in the blockchain, the algorithm can suffer from errors. Hence, multiple different types of items, transactions, and information lead to a data standardization issue and limit the true potential of the blockchain traceability method based on the algorithm.

4 Discussion

Several different methods in blockchain traceability methods involve the modification of different parts of a system. The consensus-centric method focuses on the roles of different nodes within the system. Therefore, it is implemented within the blockchain

network itself. This technique can be suitable for traceability methods where a transaction of information is uniform, and trustfulness is required. An example of such potential use can be the traceability of sensitive data of employees or financial information. It is compatible to be integrated within various systems, however, it may not be able to ‘plug and play’ from one system to another. It is possible that this method must work with another blockchain method to enhance its usability options.

The use of the public blockchain domain with the support of smart contracts is another viable method for agricultural traceability. Since more additional conditions can be set to nodes, the system can be made to be more accessible for use with traceability of various products and raw materials movement within a supply chain. Furthermore, since the public domain is used, the system is more stable than a custom-made blockchain due to more nodes can participate in the proof of stake/works. Some potential examples of public blockchains include Ethereum, Polygon Matic, and Tron. The system would require a transacting fee that usually costs a particular cryptocurrency acting as the ‘gas fee’ (Marchesi, Marchesi, Destefanis, Barabino, & Tigano, 2020). Since cryptocurrency fluctuates greatly over time, estimation of system operational costs can be proven challenging.

Another modification can be made to the algorithm of a blockchain itself. For this type of blockchain traceability method, proper mathematical modeling must be constructed from scratch for most use cases. It requires highly trained personnel to develop but offers a high degree of customization for its use case in data traceability. It can be extremely complex for data integration due to the modeling becoming increasingly convoluted. For instance, for growing only one type of crop, not only the seeds and fertilizers need to be considered. A specific type of fertilizer must also be considered for maintaining the crops. Furthermore, other parameters such as humidity-adjusting water sprinklers, soil pH maintenance, and crop movement via harvesting and logistics can add up to form an extremely complex system. Bugs and errors can be very difficult to remedy for a sophisticated system. Therefore, it can be unfavorable for the system to be easily marketed for industries to be used in their data traceability needs.

A potential gap can be identified based on the literature on data traceability of various usage scenarios. The implementation of blockchain technology for data traceability is still seldom compared to other more understood data traceability methods such as RFIDs, IoTs, and AI. Within the context of blockchain itself, several methods have been proposed for various usage scenarios, but they possess weaknesses. The most commonly found weakness in blockchain technology data traceability is the scalability, implementation, and data standardization issues. These weaknesses show a potential gap for the development of a more refined blockchain data traceability method.

5 Conclusions

This review paper presents an overview of the use of blockchain technology for data traceability in various industries. Through a thorough analysis of traditional traceability methods, the limitations of existing approaches are identified, emphasizing the need for more robust and efficient solutions. The review highlights the potential of artificial intelligence and the Internet of Things as complementary technologies to enhance

data traceability. However, challenges such as the complexity of AI and the vulnerabilities of IoT devices are acknowledged. In this context, blockchain technology emerges as a promising solution due to its decentralized and transparent nature. By exploring different traceability methods, a holistic understanding of the benefits and challenges associated with blockchain-based traceability systems is revealed. The evaluation of research studies showcases the diverse approaches taken, emphasizing the importance of factors such as data standardization, processing speed, and cost. The paper also underscores the significance of interdisciplinary collaboration, standardization efforts, and advanced analytics in further advancing blockchain-based traceability systems. By leveraging blockchain-based traceability, organizations can drive positive transformations in achieving transparency, accountability, and efficiency.

Acknowledgements. This research is supported by the Industrial Grants from ARB Cloud Sdn Bhd (No: ZG-2022-003), ARB Big Data Sdn. Bhd, (No: ZG-2022-006) and Universiti Kebangsaan Malaysia (UKM). High appreciation goes to the above sponsors.

References

- Cui, H., Chen, Z., Xi, Y., Chen, H., Hao, J.: IoT Data Management and Lineage Traceability: A Blockchain-based Solution. Paper presented at the 2019 IEEE/CIC International Conference on Communications Workshops in China (ICCC Workshops) (2019)
- George, R.V., Harsh, H.O., Ray, P., Babu, A.K.: Food quality traceability prototype for restaurants using blockchain and food quality data index. *J. Clean. Prod.* **240**, 118021 (2019). <https://doi.org/10.1016/j.jclepro.2019.118021>
- Hao, Z., Mao, D., Zhang, B., Zuo, M., Zhao, Z.: A novel visual analysis method of food safety risk traceability based on blockchain. *Int. J. Environ. Res. Public Health* **17**(7) (2020). <https://doi.org/10.3390/ijerph17072300>
- Jiang, N., Wang, W., Wu, J., Wang, J.: Traceable method for personal information registration based on blockchain. *IEEE Access* **8**, 52700–52712 (2020). <https://doi.org/10.1109/ACCESS.2020.2981175>
- Jiang, S., Li, Y., Wang, S., Zhao, L.: Blockchain competition: the tradeoff between platform stability and efficiency. *Eur. J. Oper. Res.* **296**(3), 1084–1097 (2022). <https://doi.org/10.1016/j.ejor.2021.05.031>
- Liao, Y., Xu, K.: Traceability system of agricultural product based on block-chain and application in tea quality safety management. *J. Phys. Conf. Ser.* **1288**(1), 012062 (2019). <https://doi.org/10.1088/1742-6596/1288/1/012062>
- Maity, M., Toloie, A., Sinha, A.K., Tiwari, M.K.: Stochastic batch dispersion model to optimize traceability and enhance transparency using Blockchain. *Comput. Ind. Eng.* **154**, 107134 (2021). <https://doi.org/10.1016/j.cie.2021.107134>
- Marbough, D., et al.: Blockchain for COVID-19: Review, Opportunities and a Trusted Tracking System (2020)
- Marchesi, L., Marchesi, M., Destefanis, G., Barabino, G., Tigano, D.: Design patterns for gas optimization in ethereum. Paper presented at the 2020 IEEE International Workshop on Blockchain Oriented Software Engineering (IWBOSE) (2020)
- Nyaletey, E., Parizi, R. M., Zhang, Q., Choo, K.K.R.: BlockIPFS - blockchain-enabled interplanetary file system for forensic and trusted data traceability. Paper presented at the 2019 IEEE International Conference on Blockchain (Blockchain) (2019)

- Qian, J.-P., Yang, X.-T., Wu, X.-M., Zhao, L., Fan, B.-L., Xing, B.: A traceability system incorporating 2D barcode and RFID technology for wheat flour mills. *Comput. Electron. Agric.* **89**, 76–85 (2012). <https://doi.org/10.1016/j.compag.2012.08.004>
- Qiao, R., Zhu, S., Wang, Q., Qin, J.: Optimization of dynamic data traceability mechanism in Internet of Things based on consortium blockchain. *Int. J. Distrib. Sens. Netw.* **14**(12), 1550147718819072 (2018). <https://doi.org/10.1177/1550147718819072>
- Tarjan, L., Šenk, I., Tegeltija, S., Stankovski, S., Ostojic, G.: A readability analysis for QR code application in a traceability system. *Comput. Electron. Agric.* **109**, 1–11 (2014). <https://doi.org/10.1016/j.compag.2014.08.015>
- Zhang, C., Lu, Y.: Study on artificial intelligence: The state of the art and future prospects. *J. Ind. Inf. Integr.* **23**, 100224 (2021). <https://doi.org/10.1016/j.jii.2021.100224>
- Zhu, P., Hu, J., Zhang, Y., Li, X.: A blockchain based solution for medication anti-counterfeiting and traceability. *IEEE Access* **8**, 184256–184272 (2020). <https://doi.org/10.1109/ACCESS.2020.3029196>